

abstract

COMPUTER SCIENCE/DISCRETE MATH II

Topic:

Speaker:

Affiliation:

Date:

Time/Room:

An important task in systematic biology is the reconstruction of phylogenies: The evolutionary history of a family of organisms, represented by an evolutionary tree, needs to be reconstructed from the genetic data of the extant species, which correspond to the leaves of the tree. The evolutionary process is described by a Markov process on the evolutionary tree, and the DNA sequences of the extant species are samples from this Markov process at the leaves. The task is to reconstruct the tree given the samples.

There has been a lot of work on this problem in biology, statistics and computer science. I will survey this work and conclude with an algorithm which is information-theoretically optimal, in the sense that it reconstructs a phylogeny using as little information (measured in DNA sequence length) at the leaves as possible. The algorithm uses ideas from statistical physics and provides an interesting connection of the phylogenetic reconstruction problem to the reconstruction problem for spin glasses on trees, an important problem in statistical physics and probability theory.

(based on joint work with Elchanan Mossel and Sebastien Roch)